Discovering Metamorphic Relations from OpenSWMM

Xuanyi Lin[1] (*linx7@mail.uc.edu*), Michelle Simon[2], Zedong Peng[1], Nan Niu[1]

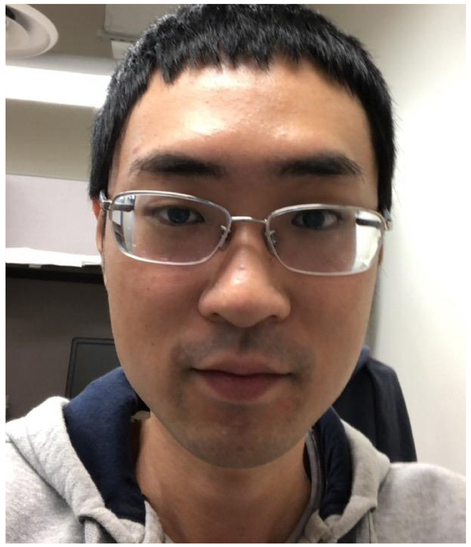
University of Cincinnati | *Department of* Electrical Engineering & Computer Science

United States Environmental Protection Agency | *Water Infrastructure Division*

February 25, 2020

Acknowledgments

University of CINCINNATI



Xuanyi Lin

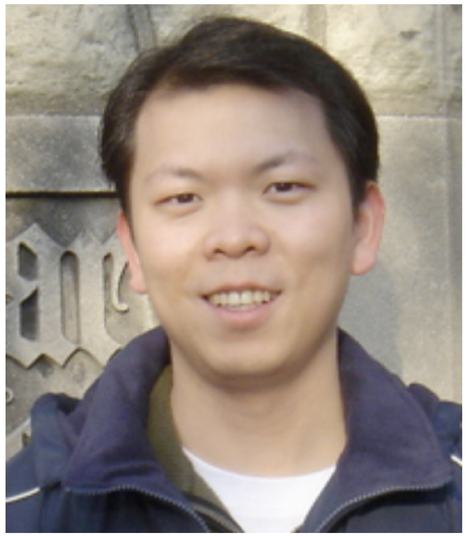
PhD Student
University of Cincinnati

Michelle Simon

Chemical Engineer
US EPA

Zedong Peng

PhD Student
University of Cincinnati



Nan Niu

Associate Professor
University of Cincinnati

USEPA Disclaimer

University of CINCINNATI

The U.S. Environmental Protection Agency, through its Office of Research and Development, funded, managed and collaborated in the research described herein. This presentation has been subjected to the Agency's administrative review and has been approved for external publication. Any opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Agency, therefore, no official endorsement should be inferred. Any mention of trade names or commercial products does not constitute endorsement or recommendation for use.

USEPA and ORISE Support

University of CINCINNATI

Xuanyi Lin was supported in part by an appointment to the Postgraduate Research Program at the (Laboratory Office of Research and Development, National Risk Management Research Laboratory) administered by the Oak Ridge Institute for Science and Education through Interagency Agreement No. (DW-8992433001) between the U.S. Department of Energy and the U.S. Environmental Protection Agency.

Outline

University of CINCINNATI

- Motivation
- Overview of Metamorphic Testing
- Metamorphic Relations Discovery from OpenSWMM
- Conclusion

Storm Water Management Model (SWMM)

University of CINCINNATI

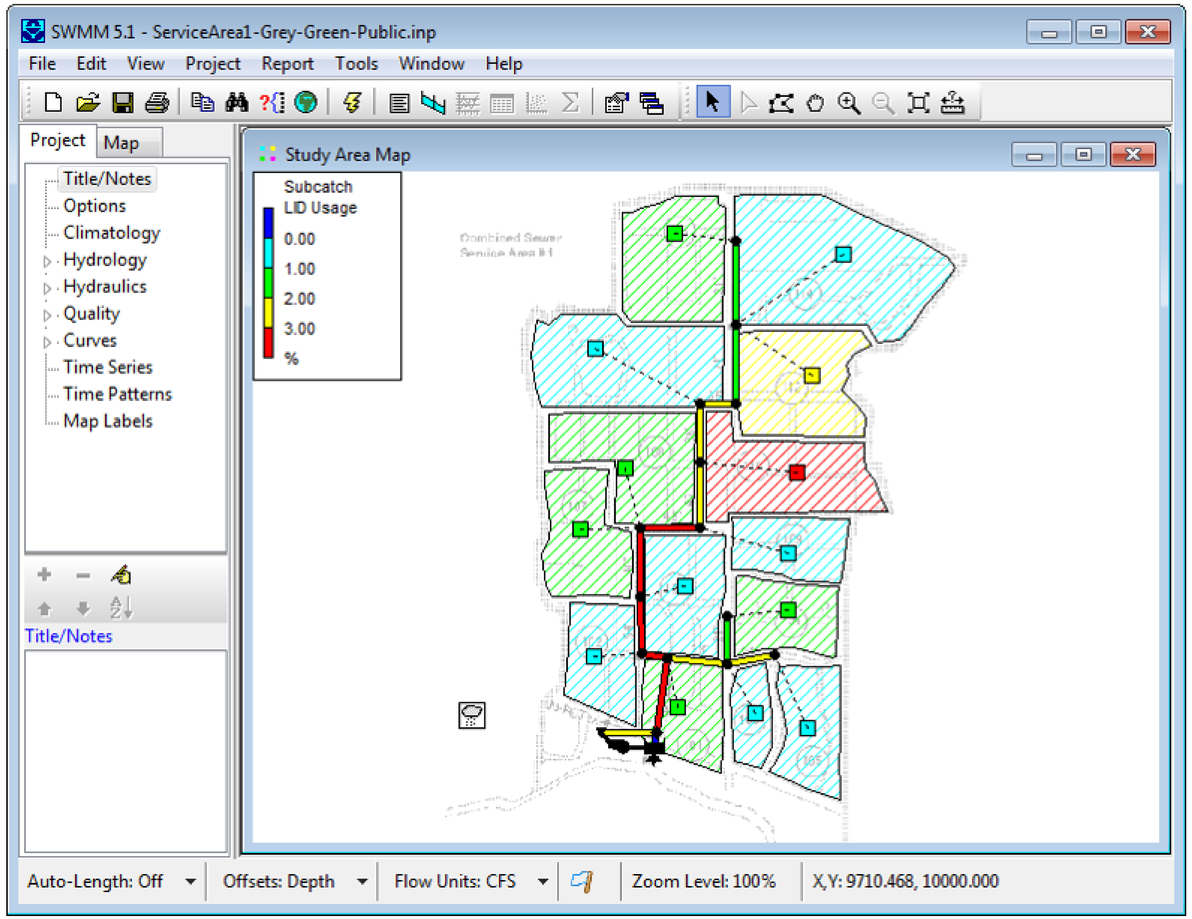
SWMM

- Used throughout the world for planning, analysis, and design related to stormwater runoff, sewers, and other drainage systems.

SWMM 5.1.13

United States Environmental Protection Agency

https://www.epa.gov/water-research/storm-water-management-model-swmm



Storm Water Management Model (SWMM)

University of CINCINNATI

SWMM

- Thousands of studies worldwide have been carried out by using SWMM.
- 21 versions of SWMM have been released since 1971.

Modelling Stormwater Pollutant Reduction with LID Scenarios
Espoo, Finland

Assessing the significance of evapotranspiration in green roof modeling
Portland, Oregon, USA

Runoff quantity response of three urban mountain parks
Chongqing, China

Partial statistics on using SWMM for research in 2018 based on Google Scholar

Outline

University of CINCINNATI

- Motivation
- Overview of Metamorphic Testing
- Metamorphic Relations Discovery from OpenSWMM
- Conclusion

Oracle & Oracle Problem

University of CINCINNATI

Oracle

- The mechanism for checking whether the program under test produces the expected output in different test cases.

Oracle problem

- Determining the correct output for a given input.

Has oracle

$1+1=2$

Do not have oracle

sine(x), cosine(x), log(x), ...

Oracle & Oracle Problem

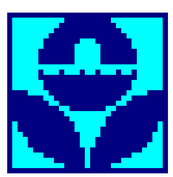
University of CINCINNATI

Oracles may not be available

- Find answers previously unknown

The impact of the modifications in nuclear weapon simulation

- Approximations involved in simulation models



rounded ← - - - - - - - - - *Parameter1= 1.23456789...*

- Do not have complete historical data

Objective assessments

Metamorphic Testing

University of CINCINNATI

Published as Technical Report HKUST-CS98-01, Department of Computer Science
The Hong Kong University of Science and Technology, 1998.

Metamorphic Testing:

A New Approach for Generating Next Test Cases †

T. Y. Chen S. C. Cheung S. M. Yiu

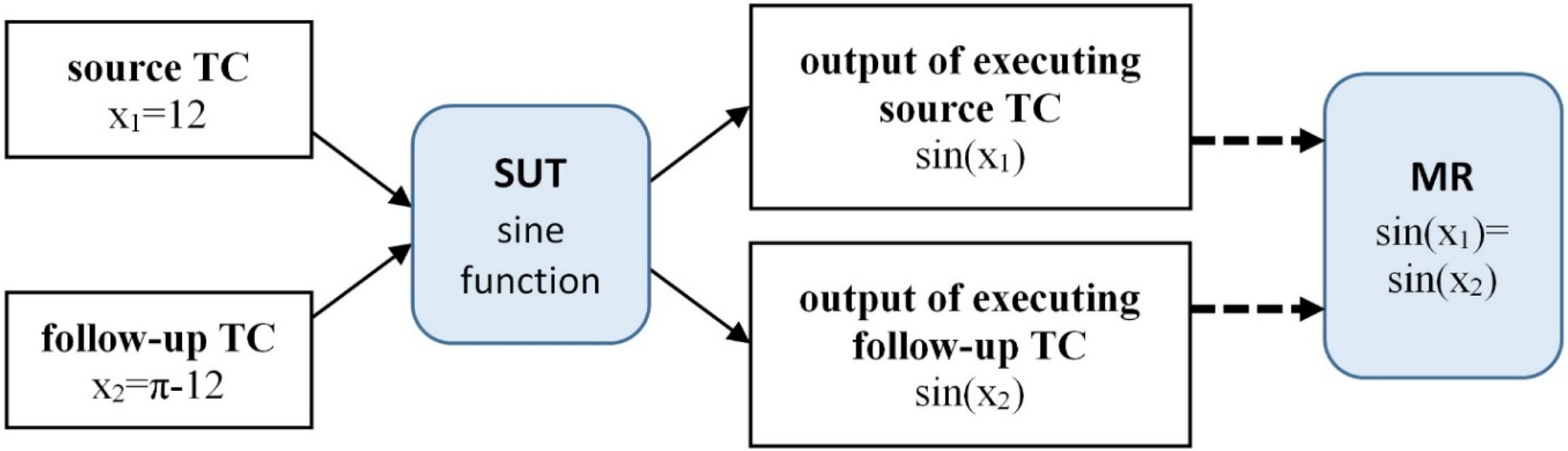
Metamorphic testing

- A property-based software testing technique to overcome the oracle problem.

Metamorphic relation

- Necessary properties of the software under test in relation to multiple inputs and their expected outputs.

Metamorphic Testing



"TC" stands for "test case", "SUT" stands for "software under test", "MR" stands for "metamorphic relation", solid arrows represent execution, and dotted arrows represent comparison.

Metamorphic relation $\sin(x) = \sin(\pi - x)$

If two test cases—12 and (π-12)—are run one after the other as shown in the above figure, then their outputs shall be equivalent regardless of the implementation specifics.

Where can we find metamorphic relations?

How useful are metamorphic relations & metamorphic testing?

Outline

University of CINCINNATI

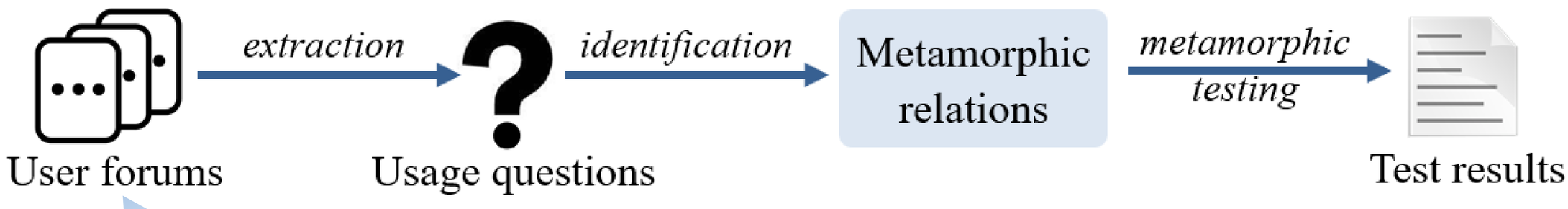
- Motivation
- Overview of Metamorphic Testing
- Metamorphic Relations Discovery from OpenSWMM
- Conclusion

Inferring Metamorphic Relations

University of CINCINNATI

User forum analysis

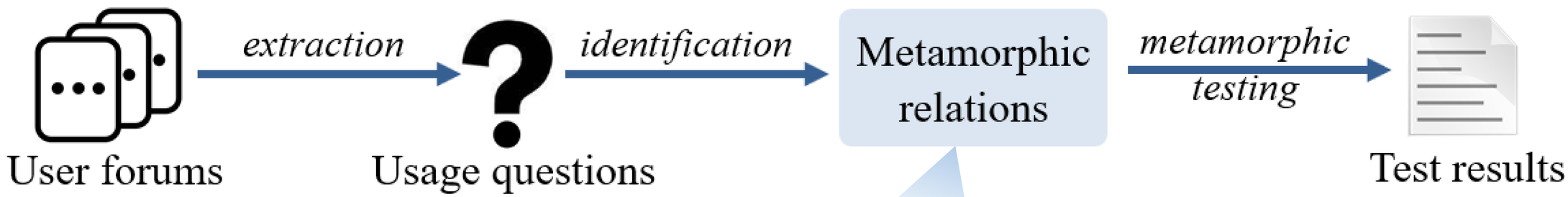
- Use the user forums as sources and explore the identification of metamorphic relations from them.



OPEN SWMM

- 25 years
- 1,818 contributors
- 15,919 posts

Metamorphic Output Relations



Hierarchical Output Relations

Category	Relation Types	Example
functional	$f(\Delta x)$	$\cos^2\left(\frac{x}{2}\right) = \frac{1}{2}\cos(x) + \frac{1}{2}$
ordinal	$<, \leq, \geq, >$	$\log(2x) > \log(x)$, $x>0$
nominal	$=, \neq, \approx$	$\sqrt{x^e} \neq \sqrt{x}$, $x>1$

Our hierarchy is rooted in the theory of measurement scales, where nominal groups data by names (e.g., “male” or “female”) and ordinal allows data to be ordered (e.g., “strongly agree” or “agree”).

University of CINCINNATI

Identify a nominal MR_{date} from the post

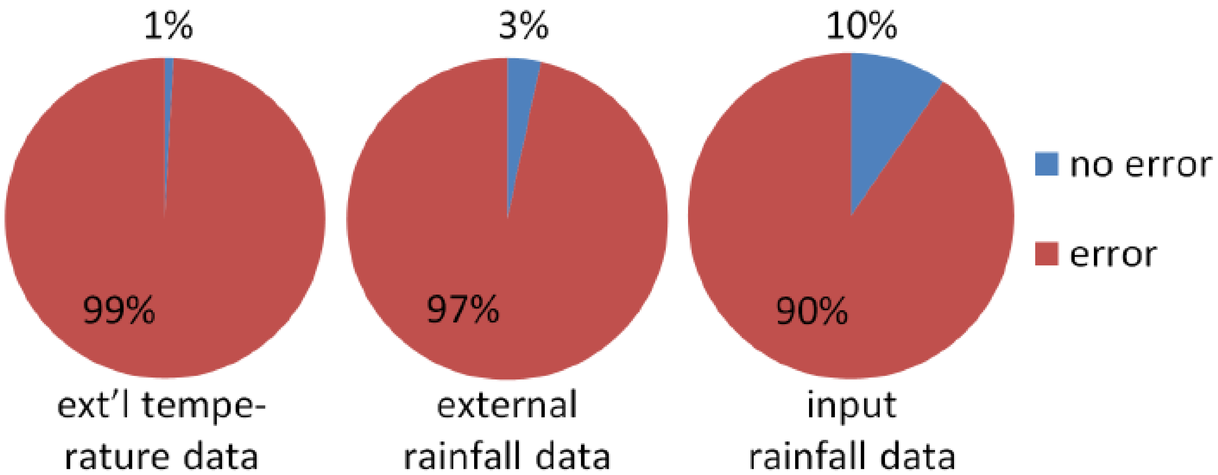
User1 08-Feb-2018

...

I tried to rotate the external rainfall data in .dat format and the error always appears: "ERROR 363: invalid data in external file used for Time Series A4_2." I have already modified the format in different ways, but the error always appears.

```
Format 1: A4_2 NOV-26-2016 00:01:00 13.7
Format 2: A4_2 11-26-2016 00:01:00 13.7
Format 3: A4_2 11-26-2016 00:01:00 13.7
Format 4: A4_2 2016 11 26 00:01:00 13.7
```

...



Testing results of MR_{date} where each pie chart is based on 480 test cases.

MR_{date}: Adjust the date format in the input data and check for errors at runtime.

Ordinal Metamorphic Relations Depict Behavioral Trend

University of CINCINNATI

Identify an ordinal MR_{thread} from the post

User3 15-Feb-2018

I am simulating 3 urban catchments (totally 15.7 Km2). There are 273,000 stormwater pipes and the period of simulation is 18 years (rainfall data). When I only run the area for 2 months period (6 minutes time step) program shows up to 150 hours for running.

...

Commenter2

...

[Finally], are you using the maximum number of threads?

User3

Many thanks for your useful comments. Yes, I have 237K pipes. I changed the threads from 1 to 4 and the simulation time has changed significantly. Now I am running for 10 years rainfall within 10 minutes.

...

MR_{thread}: Increase the number of threads and check for simulation speedup.

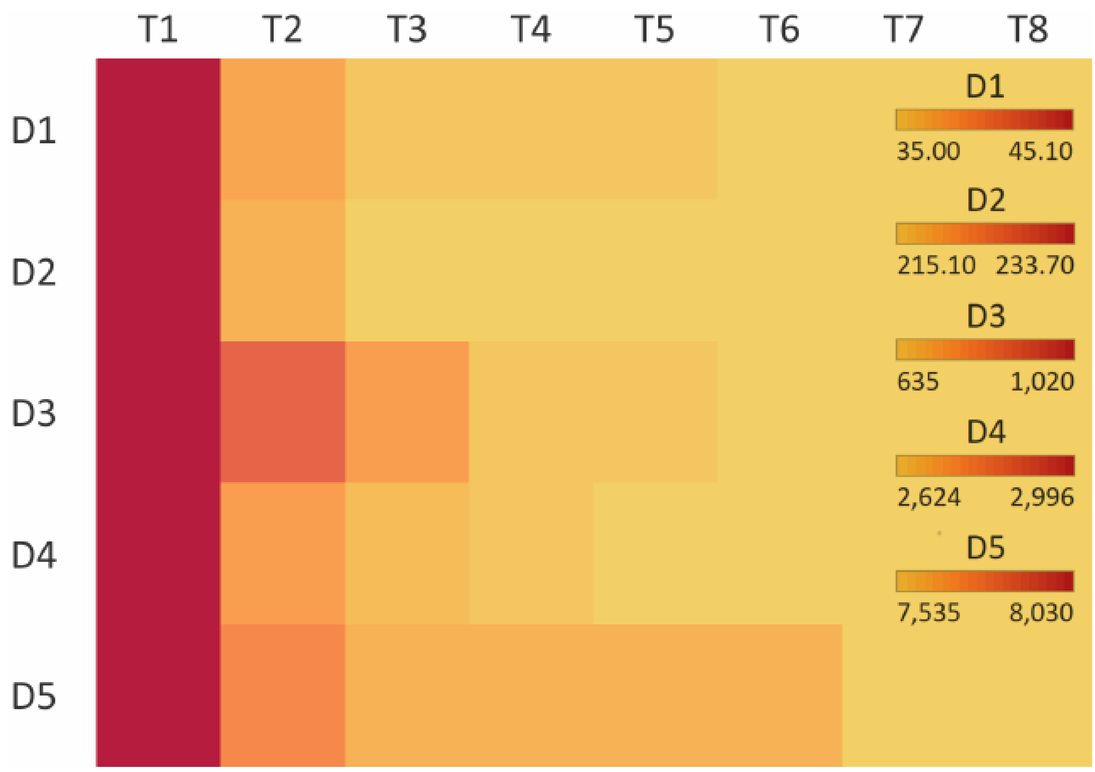
	T1	T2	T3	T4	T5	T6	T7	T8
D1	45.1	37.2	35.8	35.7	35.6	35.2	35.2	35.0
D2	233.7	218.3	215.9	215.7	215.1	215.1	215.4	215.8
D3	1020	850	741	669	660	643	640	635
D4	2996	2729	2665	2661	2624	2629	2627	2626
D5	8030	7725	7630	7630	7625	7625	7545	7535

Testing MR_{thread} on five datasets (rows) with an increasing number of threads/columns (one to eight).
SWMM execution time (seconds) for completing 100 simulations.

Ordinal Metamorphic Relations Depict Behavioral Trend

University of CINCINNATI

Identify an ordinal MR_{thread} from the post

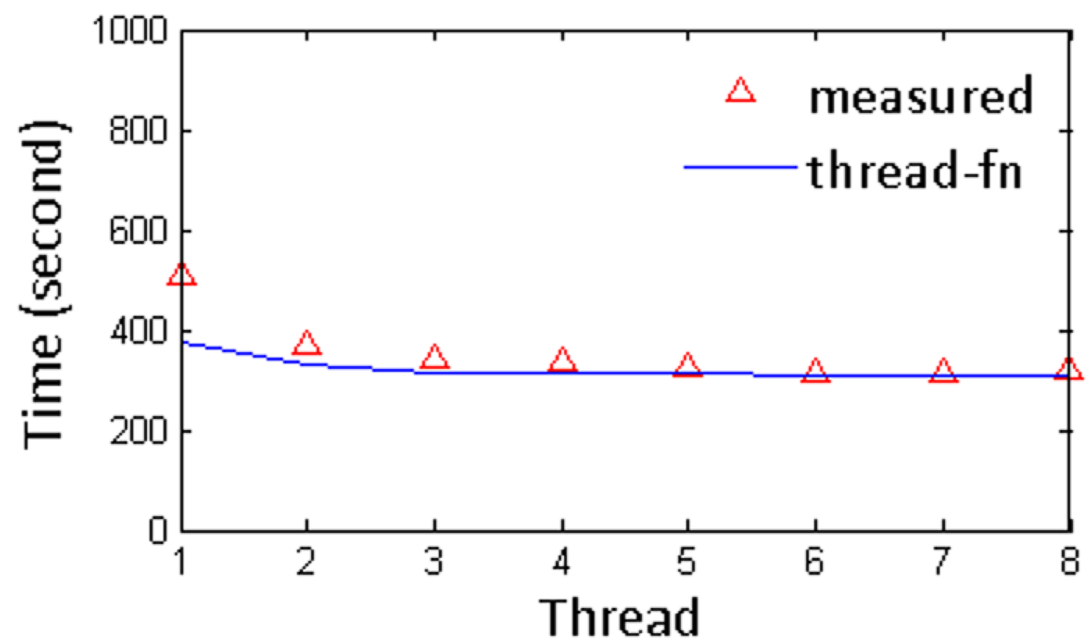


Plotting SWMM speeds in colored cells:
The lighter the cell, the faster SWMM runs.

Regression analysis in MATLAB:

$$speed=(a\times t^4+b\times t^3+c\times t^2+d\times t+e)\times C$$

C=(size+# of nodes)×I/O reporting÷500,
$a=7.57\times10^{-4}$, $b=0.017$, $c=0.129$, $d=0.437$, $e=1.620$.



Build a functional MR_{weir}

User4 08-Mar-2019

I'm modelling an open storm drain (ditch) that at one point has an abrupt change in slope, going from very steep to nearly horizontal. This is having a strange effect, and the junction at the transition seems to be functioning as a retention basin - Qin >> Qout during peak flow and then Qout > Qin. No siginficant losses and no flooding.
How can I work around this so that the flow transitions normally? ...
it doen't make sense to me.
...

Commenter3

...
Nodes in SWMM contain half of the volume of each connected link, so there is a storage effect.
Qin *dt = dV + Qout *dt.
...

MR_{Weir}: Apply different weir types at runtime and check for the quantitative prediction.

Weir Type	Cross Section Shape	Flow Formula
TRANSVERSE	Rectangular	$C_w S h^{3/2}$
SIDEFLOW	Rectangular	$C_w L h^{5/3}$
V-NOTCH	Triangular	$C_w S h^{5/2}$
TRAPEZOIDAL	Trapezoidal	$C_w L h^{3/2} + C_{ws} S h^{5/2}$
ROADWAY	Rectangular	$C_w L h^{3/2}$

C_w = weir discharge coefficient, L = weir length, S = side slope of V-NOTCH or TRAPEZOIDAL weir, h = head difference across weir, C_{ws} = discharge coefficient through sides of TRAPEZOIDAL weir.

Build a functional MR_{weir}

User4 08-Mar-2019

I'm modelling an open storm drain (ditch) that at one point has an abrupt change in slope, going from very steep to nearly horizontal. This is having

End-user debugging is difficult on those variables whose values are not directly visible.

Commenter3

...

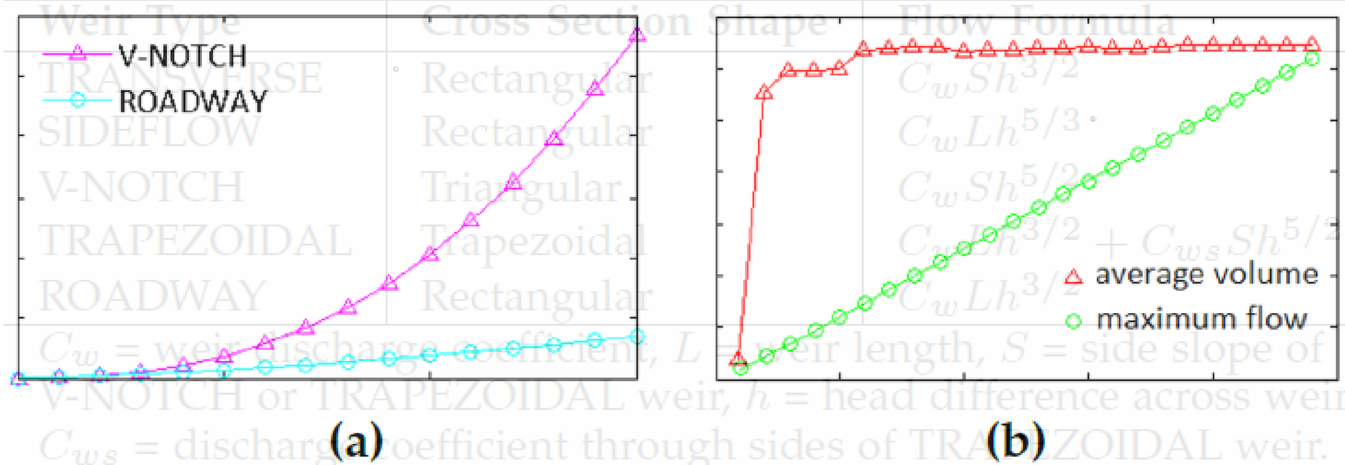
Nodes in SWMM contain half of the volume of each connected link, so there is a storage effect. Qin *dt = dV + Qout *dt.

...

MR_{Weir}: Apply different weir types at runtime and check for the quantitative prediction.

Weir Type	Cross Section Shape	Flow Formula
TRANSVERSE	Rectangular	$C_w S h^{3/2}$
SIDEFLOW	Rectangular	$C_w L h^{5/3}$
V-NOTCH	Triangular	$C_w S h^{5/2}$
TRAPEZOIDAL	Trapezoidal	$C_w L h^{3/2} + C_{ws} S h^{5/2}$
ROADWAY	Rectangular	$C_w L h^{3/2}$

C_w = weir discharge coefficient, L = weir length, S = side slope of V-NOTCH or TRAPEZOIDAL weir, h = head difference across weir, C_{ws} = discharge coefficient through sides of TRAPEZOIDAL weir.



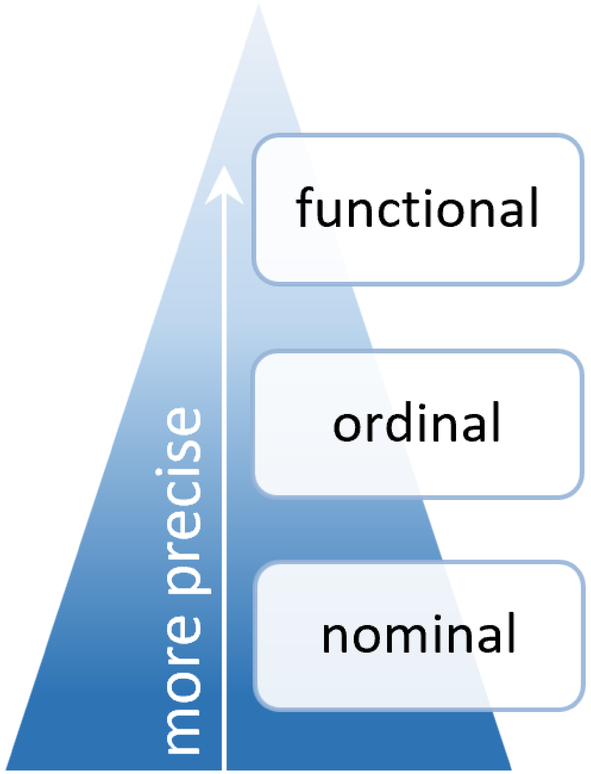
(a) Plotting V-NOTCH and ROADWAY on h;
(b) Testing MR_{weir} of V-NOTCH and ROADWAY on two SWMM output variables

Metamorphic Relation Results

University of CINCINNATI

The forums where end users post questions and exchange concerns are valuable sources to discover metamorphic relations.

MR	EXPLANATION
MR_{weir}	apply different weir types at runtime and check for the quantitative prediction
MR_{ditch}	increase inflow and check for outflow
$MR_{iteration}$	increase simulation iterations and check for reduced continuity error
MR_{date}	adjust the date format in the input data and check for errors at runtime
MR_{file}	change the format type of input data file and check for errors at runtime
MR_{flow}	select different flow routing methods and check if the output varies



Outline

University of CINCINNATI

- Motivation
- Overview of Metamorphic Testing
- Metamorphic Relations Discovery from OpenSWMM
- Conclusion

Conclusion

University of CINCINNATI

- Introduce the method of inferring metamorphic relations from Open SWMM forum.
- Demonstrate the usefulness of metamorphic relations identified from user's perspective.
- Suggest a hierarchy of nominal, ordinal, and functional MRs.

Questions

University of CINCINNATI

Thanks!

UNIVERSITY OF Cincinnati

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY